

# Deterministic evolution of an asexual population under the action of beneficial and deleterious mutations on additive fitness landscapes

Kavita Jain and Sona John

*Theoretical Sciences Unit, Jawaharlal Nehru Centre for Advanced Scientific Research,  
Jakkur P.O., Bangalore 560064, India*

---

## Abstract

We study a continuous time model for the frequency distribution of an infinitely large asexual population in which both beneficial and deleterious mutations occur and fitness is additive. When beneficial mutations are ignored, the exact solution for the frequency distribution is known to be a Poisson distribution. Here we include beneficial mutations and obtain exact expressions for the frequency distribution at all times using an eigenfunction expansion method. We find that the stationary distribution is non-Poissonian and related to the Bessel function of the first kind. We also provide suitable approximations for the stationary distribution and the time to relax to the steady state. Our exact results, especially at mutation-selection equilibrium, can be useful in developing semi-deterministic approaches to understand stochastic evolution.

*Key words:* mutation-selection balance, exact results, beneficial mutations

---

## 1. Introduction

Real populations are finite and evolve stochastically. However theoretical investigations of infinitely large populations subject to various evolutionary forces have been useful in developing diffusion theory, theory of branching process and semi-deterministic approaches to describe the evolution of finite populations [11]. Therefore, a detailed analytical study of deterministic models constitutes an important step towards an understanding of more complex and realistic situations. In the simplest scenario, one may consider an asexual population of infinite size under mutation and selection. When the mutation

rates depend linearly on the number of deleterious mutations in a sequence and the fitness of the genetic sequence is additive, a complete solution of the genotypic frequency distribution is known exactly [26, 13] and has been utilised, for example, in modeling codon usage bias [16, 3, 18].

Another model in which the fitnesses are non-epistatic but the rate of deleterious and beneficial mutations is independent of the fitness of the sequence has recently appeared in various contexts such as adaptive evolution [23, 14, 5, 21, 19], evolution of sex [9] and evolution of mutation rates [12]. For this mutation scheme, when beneficial mutations are ignored, the exact solution in the stationary state and for the dynamics of the population fraction is known; in particular, at mutation-selection balance, the frequency is Poisson-distributed with a mean given by the ratio of the deleterious mutation rate to selection coefficient [15, 10, 17, 7]. When beneficial mutations are also allowed, the short time dynamics of the frequency distribution have been quite well studied [19] and some approximate results at mutation-selection equilibrium were obtained recently [12]. Here we study this model in detail and find an exact expression for the frequency distribution at all times.

In the following section, we define the model, discuss some limiting cases and explain the relation of our work to the existing literature. We then proceed to find an exact solution of the stationary state as well as the dynamics using an eigenfunction expansion method in Sec. 3. Besides the exact results, we also provide accurate approximations for the frequency distribution when the selection coefficient is larger or smaller than the mutation rates. Sections 4 and 5, respectively, deal with these approximations in the stationary state and for the dynamics of the frequency distribution. A discussion of the results follows in the concluding section.

## 2. Model

We consider an infinitely large asexual population of infinitely long diallelic sequences evolving in continuous time. All individuals carrying  $k \geq 0$  deleterious mutations relative to the fittest individual are assumed to have the same (Malthusian) fitness  $w_k = -sk, s \geq 0$  and said to belong to the fitness class  $k$ . We also assume a single-step mutation scheme in which a deleterious (beneficial) mutation increases (decreases) the fitness class by one and occurs at rate  $U_d$  ( $U_b$ ), but mutations to other classes are ignored. Then the population fraction  $x_k(t)$  in the  $k$ th fitness class at time  $t$  obeys

the following differential-difference equations:

$$\dot{x}_0(t) = U_b x_1(t) - U_d x_0(t) + s \mathcal{C}_1(t) x_0(t) , \quad (1a)$$

$$\dot{x}_k(t) = U_b x_{k+1}(t) + U_d x_{k-1}(t) - U x_k(t) - s(k - \mathcal{C}_1(t)) x_k(t) , \quad k \geq 1 . \quad (1b)$$

In the above equations,  $U = U_d + U_b$  is the total mutation rate and  $\mathcal{C}_1(t) = \sum_{k=0}^{\infty} k x_k(t)$  is the average number of deleterious mutations in the population at time  $t$ . In the stationary state where the LHS of (1a) and (1b) is zero, we will denote the steady state fraction by  $x_k$ .

It is easy to verify that the above set of equations respect the normalisation condition,

$$\sum_{k=0}^{\infty} x_k(t) = 1 , \quad t \geq 0 . \quad (2)$$

We also have the boundary condition

$$x_k(t) \xrightarrow{k \rightarrow \infty} 0 , \quad (3)$$

which ensures that the total fraction remains finite at all times. To complete the definition of the model, we also need to specify the initial condition  $x_k(0)$  for all  $k$ . The analysis in Sec.s 3 and 4 holds for arbitrary initial conditions but in Sec. 5, we will assume that the population is initially localised in the fitness class  $k^{(0)}$ .

The time evolution equations above for the frequency  $x_k(t)$  are defined for  $k \geq 0$  and therefore the maximum fitness is zero. However, (1b) has been used without an upper bound on fitness to describe the adaptation dynamics [23, 14, 5, 21, 2, 19]. The latter is a reasonable model at short times for a population initially localised in a fitness class with many deleterious mutations since the frequency in the fitness classes close to the fittest one can then be neglected. For this model, several works [23, 14, 5, 19] have shown that it does not have a traveling wave solution, and either a lower cutoff on the frequency modeling a finite population size [23] or discrete time dynamics [19] are required to obtain it. In [21], although a cutoff for the high-fitness edge is imposed in the deterministic model to account for the finite size of the population, this work also assumes a traveling wave solution for the continuous time model (see their (4) and (5)). However, as in [23, 14, 5, 19], our analysis of short time dynamics described in Sec. 5.1 also does not support a traveling wave behavior.

Although dynamics can be studied on an infinite line, a stationary state does not exist if the fitness is not bounded above. To see this, consider the steady state of (1b) by setting the LHS to be zero. Since the frequency in any fitness class must not be negative and the first two terms on the RHS of (1b) are positive, their contribution can be balanced if

$$k > \mathcal{C}_1 - U/s = k_* , \quad (4)$$

thus leading to a maximum fitness corresponding to  $-sk_*$ . A previous analysis of (1b) with unbounded fitnesses finds a negative frequency distribution in the stationary state and claims that “in the deterministic limit there is no true stationary state for arbitrary [beneficial mutation rate]...” (p. 1313, [9]). However as discussed above, the loss of positivity is simply a consequence of the lack of upper bound on the fitness and in Sec. 4, we will show that the model defined by (1a) and (1b) has a nontrivial steady state. We also mention that if we set the maximum fitness to  $-sk_*$  instead of zero, the frequency  $x_m = 0, m < k_*$  and  $x_{m+k}, m \geq k_*$  is given by the solution  $x_k$  of (1a) and (1b) [10].

Equations (1a) and (1b) are mathematically nontrivial for two reasons: first, they are nonlinear in the fractions  $x_k(t)$  due to the last (selection) term on the RHS and second, they are second order difference equations in  $k$  when both mutation rates are nonzero.

(i) When the beneficial mutations are absent ( $U_b = 0$ ), in the stationary state, the boundary equation (1a) immediately yields the average number of deleterious mutations,  $\mathcal{C}_1 = U_d/s$ . This result is very helpful since it renders (1b) to be linear in the frequencies and we quickly arrive at the following well known result [15, 10]:

$$x_k = \frac{e^{-U_d/s}}{k!} \left( \frac{U_d}{s} \right)^k \quad [U_b = 0] . \quad (5)$$

The time-dependent frequency has also been obtained using a generating function method and shown to be Poisson-distributed [7].

(ii) When the deleterious mutations are absent ( $U_d = 0$ ), the stationary state is trivial ( $x_k = \delta_{k,0}$ ). But the short time dynamics can be obtained by extending the method of [7] as described in Sec. 5.1 (also, see [14, 5, 19]).

(iii) In the neutral case ( $s = 0$ ), the nonlinear term on the RHS of (1a) and (1b) vanishes. The stationary state frequency is then easily found to be

$$x_k = \left( 1 - \frac{U_d}{U_b} \right) \left( \frac{U_d}{U_b} \right)^k \quad [s = 0, U_d < U_b] . \quad (6)$$

	$U_b = 0$	$U_d = 0$	$s = 0$	all nonzero
Stationary state	[15, 10]	trivial	none	Equation (27)
Dynamics	[7]	[14, 5, 19]	Equation (A6)	Equation (26)

Table 1: Summary of the results for the deterministic model defined by (1a) and (1b) where  $U_d$  and  $U_b$ , respectively, denote deleterious and beneficial mutation rate and  $s$  is the selection coefficient. In all the cases except when deleterious mutations are absent, it is assumed that  $U_b < U_d$ .

The condition  $U_d < U_b$  arises due to the boundary condition (3). However, in the parameter regime where  $U_b < U_d$ , the neutral population does not reach a steady state. In Appendix A1, we give the exact solution of the neutral dynamics in this parameter regime.

A brief summary of the results in the stationary state and for the dynamics is given in Table 1.

### 3. Exact solution of the population frequency by eigenfunction expansion method

We now proceed to find the population fraction  $x_k(t)$  when all the three parameters, viz.,  $s, U_b, U_d$  are nonzero. In the following discussion, we assume that the mutation rate  $U_d > U_b$  as in biologically realistic situations [20]. Since the equations (1a) and (1b) are nonlinear in the fractions  $x_k(t)$ , we work with the unnormalised variables defined as [22, 11]

$$z_k(t) = x_k(t) e^{-s \int_0^t dt' C_1(t')}, \quad (7)$$

which obey the following *linear* equations:

$$\dot{z}_0(t) = U_b z_1(t) - U_d z_0(t), \quad (8a)$$

$$\dot{z}_k(t) = U_b z_{k+1}(t) + U_d z_{k-1}(t) - U z_k(t) - s k z_k(t), \quad k \geq 1. \quad (8b)$$

Summing over  $k$  on both sides of (7) and using the normalisation condition (2), we obtain the following relationship between the average  $C_1(t)$  and the unnormalised frequencies  $z_k(t)$ :

$$\sum_{k=0}^{\infty} z_k(t) = e^{-s \int_0^t dt' C_1(t')}. \quad (9)$$

Using this in (7), we immediately obtain

$$x_k(t) = \frac{z_k(t)}{\sum_{m=0}^{\infty} z_m(t)} . \quad (10)$$

It is convenient to further define (p. 139, [24])

$$y_k(t) = \left( \frac{U_b}{U_d} \right)^{k/2} e^{(\sqrt{U_d} - \sqrt{U_b})^2 t} z_k(t) , \quad (11)$$

$$\tau = t \sqrt{U_b U_d} , \quad (12)$$

$$\gamma = 2 - \sqrt{U_b/U_d} , \quad (13)$$

$$S = \sqrt{s/U_b} \cdot \sqrt{s/U_d} . \quad (14)$$

In terms of these variables, we have

$$\frac{\partial y_0(\tau)}{\partial \tau} = y_1(\tau) - \gamma y_0(\tau) , \quad (15a)$$

$$\frac{\partial y_k(\tau)}{\partial \tau} = y_{k+1}(\tau) + y_{k-1}(\tau) - (2 + Sk)y_k(\tau) , \quad k \geq 1 . \quad (15b)$$

The above set of equations involving two independent variables, viz., space and time can be solved by the eigenfunction expansion method (Chapter 5 and 6, [24]). Since the differential operator  $\partial/\partial\tau$  has eigenfunctions  $e^{-\lambda\tau}$  with eigenvalues  $-\lambda$ , on expanding  $y_k(\tau)$  as a linear combination of these eigenfunctions as

$$y_k(\tau) = \sum_{\lambda} c_{\lambda} e^{-\lambda\tau} \phi_k(\lambda) , \quad k \geq 0 , \quad (16)$$

we obtain difference equations in one independent variable:

$$\phi_1 - (\gamma - \lambda)\phi_0 = 0 , \quad (17a)$$

$$\phi_{k+1} + \phi_{k-1} - (2 + Sk - \lambda)\phi_k = 0 , \quad k \geq 1 . \quad (17b)$$

Equation (17b) is an eigenvalue equation for a real symmetric matrix with eigenfunction  $\phi_k$  and eigenvalue  $-\lambda$ . For such a matrix, it is possible to find a complete set of eigenvectors [4]. Moreover, the eigenvalues are real and the eigenfunctions corresponding to different eigenvalues are orthogonal and can be normalised to unity:

$$\sum_{k=0}^{\infty} \phi_k(\lambda) \phi_k(\lambda') = \delta_{\lambda, \lambda'} . \quad (18)$$

The eigenvalues are determined by the boundary condition (17a) as explained below. The constants  $c_\lambda$ 's in (16) can be found using the initial condition and are given by

$$c_\lambda = \sum_{m=0}^{\infty} \phi_m(\lambda) y_m(0) , \quad (19)$$

$$= \sum_{m=0}^{\infty} \phi_m(\lambda) \left( \frac{U_b}{U_d} \right)^{m/2} x_m(0) . \quad (20)$$

This can be seen by using (16) at  $t = 0$  and using the orthonormality condition (18).

Our remaining task now is to find the eigenfunctions  $\phi_k$ . We remark that if the fitness class  $k$  is treated as a continuous variable, (17b) reduces to a time-independent Schrödinger equation for a particle in a linear potential for which the eigenfunctions are known to be Airy function (and plane wave when  $S$  is zero) [8]. Here we are interested in finding the eigenfunctions in discrete fitness space with (Robin) boundary condition (17a). In the neutral case ( $S = 0$ ), exact eigenfunctions and frequency  $y_k(t)$  are obtained in Appendix A1. When the parameter  $S$  is nonzero, the solution of (17b) is a linear combination of the Bessel function of first and second kind with order  $\nu$  and argument  $z$  that are denoted by  $J_\nu(z)$  and  $Y_\nu(z)$ , respectively [6] (also, see Appendix A2):

$$\phi_k(\lambda) = A(\lambda) J_{k+\frac{2-\lambda}{S}} \left( \frac{2}{S} \right) + A'(\lambda) Y_{k+\frac{2-\lambda}{S}} \left( \frac{2}{S} \right) , \quad k \geq 0 . \quad (21)$$

It is easy to check that (21) satisfies the eigenvalue equation (17b) using the recurrence relation for the Bessel function  $\mathcal{K}_\nu(z)$  given by (9.1.27, [1])

$$\mathcal{K}_{\nu-1}(z) + \mathcal{K}_{\nu+1}(z) = \frac{2\nu}{z} \mathcal{K}_\nu(z) , \quad (22)$$

where  $\mathcal{K}$  denotes  $J, Y$ . To find the constants  $A, A'$ , we invoke the boundary condition (3) that the frequency  $x_k(t) \rightarrow 0$  as the fitness class  $k \rightarrow \infty$ . From (7) and (11), it follows that  $z_k(t)$  and  $y_k(t)$  also obey this boundary condition. Using this large  $k$  behavior, we find that the coefficient  $A'$  is zero since  $Y_\nu(z)$  diverges for large  $\nu$  (9.3.1, [1]). Then using the orthonormality condition (18), we find that

$$A^2(\lambda) = \frac{1}{\sum_{k=0}^{\infty} J_{k+\frac{2-\lambda}{S}}^2 \left( \frac{2}{S} \right)} . \quad (23)$$

The eigenvalues are determined using (21) in the boundary condition (17a) at  $k = 0$  and satisfy

$$J_{1+\frac{2-\lambda}{S}}\left(\frac{2}{S}\right) - (\gamma - \lambda)J_{\frac{2-\lambda}{S}}\left(\frac{2}{S}\right) = 0 . \quad (24)$$

Putting all the pieces together, we finally obtain

$$x_k(t) \propto \left(\frac{U_d}{U_b}\right)^{k/2} \sum_{\lambda} c_{\lambda} A(\lambda) J_{k+\frac{2-\lambda}{S}}\left(\frac{2}{S}\right) e^{-\lambda\sqrt{U_b U_d}t} , \quad (25)$$

where  $c_{\lambda}$  and  $A(\lambda)$  are, respectively, given by (20) and (23), the eigenvalues by (24) and the proportionality constant is determined by the normalisation condition (2). If  $\lambda_{\alpha}, \alpha \geq 0$  denotes the  $(\alpha + 1)$ th minimum eigenvalue, the above equation can be rewritten as

$$x_k(t) = \frac{\left(\frac{U_d}{U_b}\right)^{k/2} \sum_{\alpha=0}^{\infty} c_{\lambda_{\alpha}} A(\lambda_{\alpha}) J_{k+\frac{2-\lambda_{\alpha}}{S}}\left(\frac{2}{S}\right) e^{-(\lambda_{\alpha}-\lambda_0)\sqrt{U_b U_d}t}}{\sum_{m=0}^{\infty} \left(\frac{U_d}{U_b}\right)^{m/2} \sum_{\alpha=0}^{\infty} c_{\lambda_{\alpha}} A(\lambda_{\alpha}) J_{m+\frac{2-\lambda_{\alpha}}{S}}\left(\frac{2}{S}\right) e^{-(\lambda_{\alpha}-\lambda_0)\sqrt{U_b U_d}t}} . \quad (26)$$

This result can be verified by plugging it in (1a) and (1b) and using the relationship (9) between the normalisation constant and the mean.

#### 4. Stationary state frequency

To obtain the steady state, we take the limit  $t \rightarrow \infty$  in (26) and find that only the minimum eigenvalue  $\lambda_0$  contributes to the sum over the eigenvalues and the result is independent of the initial condition. We thus obtain the *exact* stationary state frequency for an infinitely large population evolving under the joint action of deleterious and beneficial mutations and non-epistatic selection to be

$$x_k = \frac{\left(\frac{U_d}{U_b}\right)^{k/2} J_{k+\frac{2-\lambda_0}{S}}\left(\frac{2}{S}\right)}{\sum_{m=0}^{\infty} \left(\frac{U_d}{U_b}\right)^{m/2} J_{m+\frac{2-\lambda_0}{S}}\left(\frac{2}{S}\right)} , \quad (27)$$

where  $\lambda_0$  is the minimum eigenvalue determined from (24). Before proceeding further, we note that the Bessel function  $J_{\nu}(z)$  is an oscillatory function in both  $\nu$  and  $z$ . However, since the eigenfunction corresponding to the minimum eigenvalue for a real symmetric matrix with homogeneous boundary



condition cannot have zeros (p. 452, [4]), the solution (27) satisfying (17a) and (17b) is guaranteed to be positive for all  $k \geq 0$ .

Using the above solution in (1a), we find that the average number of deleterious mutations in the steady state is given exactly by

$$\mathcal{C}_1 = \frac{U_d}{s} - \frac{\gamma - \lambda_0}{S} \quad (28)$$

$$= \frac{U}{s} - \frac{2 - \lambda_0}{S}. \quad (29)$$

As Fig. 1 shows, the minimum eigenvalue  $\lambda_0$  initially increases with  $S$  and approaches a constant asymptotically. Taking  $S \rightarrow \infty$  in (24) and using that  $J_0(0) = 1, J_1(0) = 0$  [1], we find that  $\lambda_0 \rightarrow \gamma$  (also, see (31) below). Then, from (28), it follows that beneficial mutations decrease the average number of deleterious mutations as also expected intuitively. Moreover, using the inequalities  $\lambda_0 \leq \gamma \leq 2$  in (29), it is easily checked that the condition (4) for the existence of the stationary state is satisfied.

The higher order cumulants such as variance and skewness can be found using a cumulant generating function as detailed in Appendix A3. Alternatively, on multiplying both sides of (1b) in the steady state by  $k$  and summing over  $k$ , we find the stationary state variance  $\mathcal{C}_2 = \overline{k^2} - \overline{k}^2$  to be

$$\mathcal{C}_2 = \frac{U_d}{s} - \frac{U_b}{s}(1 - x_0), \quad (30)$$

which shows that beneficial mutations decrease the width of the distribution also. Furthermore, as the inset of Fig. 2 shows, the variance to mean ratio is greater than one and therefore the frequency distribution is non-Poissonian when  $U_b$  is nonzero.

To obtain some insight into the behavior of the equilibrium frequency given by (27) above, we now consider two parameter regimes where the ratio  $S$  of the selection coefficient to the mutation rates is large or small relative to one.

#### 4.1. When the parameter $S$ is large

The parameter  $S \gg 1$  when (i)  $U_b < s$  and (ii) either  $U_d < s$  or  $s < U_d < s^2/U_b$ . When  $U_b = 0$ , as (5) shows, the average number of deleterious mutations in the stationary state equals  $U_d/s$ . Therefore, when  $U_b$  is turned on, we expect that beneficial mutations do not have a significant effect when

$U_d < s$  but they can decrease the mean  $\mathcal{C}_1$  substantially when  $U_d > s$ . The analysis given below is in agreement with these expectations.

As described in Appendix A4 and shown in Fig. 1, the minimum eigenvalue  $\lambda_0$  when  $S$  is large is given by

$$\lambda_0 \approx \gamma - S^{-1} . \quad (31)$$

On plugging (31) in (27), we obtain

$$x_k \propto \left( \frac{U_d}{U_b} \right)^{k/2} J_{k + \frac{U_b}{s}(1 + \frac{U_d}{s})} \left( \frac{2}{S} \right) , \quad (32)$$

which, on using the series representation (A17) of Bessel function, yields

$$x_k \propto \left( \frac{U_d}{s} \right)^k \sum_{m=0}^{\infty} \left( \frac{U_b U_d}{s^2} \right)^m \frac{(-1)^m}{m!(m + k + \frac{U_b}{s}(1 + \frac{U_d}{s}))!} . \quad (33)$$

We check that the above solution reduces to (5) when  $U_b = 0$ . To see the effect of beneficial mutations, it is useful to expand (33) in a power series in  $U_b/s$  as was done recently in [12] and described here briefly in Appendix A5. This discussion as also Fig. 3 show that a nonzero  $U_b$  has a significant effect when  $U_d > s$ .

Furthermore, using (31) in the exact expression (28) for the average  $\mathcal{C}_1$ , we find that [12]

$$\mathcal{C}_1 \approx \frac{U_d}{s} \left( 1 - \frac{U_b}{s} \right) . \quad (34)$$

Since  $U_b/s$  is small for large  $S$ , using (5) for the frequency  $x_0$  in (30), we find that the variance is well approximated by

$$\mathcal{C}_2 \approx \frac{U_d}{s} - \frac{U_b}{s} (1 - e^{-U_d/s}) , \quad (35)$$

$$= \begin{cases} \mathcal{C}_1 & , U_d \ll s , \\ \frac{U_d - U_b}{s} & , U_d \gg s . \end{cases} \quad (36)$$

Thus the variance is close to mean (34) when  $U_d/s \ll 1$  but larger in the opposite parameter regime. The above approximations are tested against the corresponding exact results in Fig. 2 and we see a good agreement.

#### 4.2. When the parameter $S$ is small

The parameter  $S \ll 1$  when (i)  $s < U_d$  and (ii) either  $s < U_b$  or  $U_b < s < \sqrt{U_b U_d}$ . A biologically relevant situation where  $S$  can be small arises in the case of mutators where mutation rates can be as high as  $10^{-2}$  [25]. Then for selection coefficient in the range  $10^{-4} - 10^{-3}$ , one obtains  $S \sim 0.01 - 0.1$ .

For small  $S$ , the minimum eigenvalue shown in Fig. 1 is calculated in Appendix A6 and given by

$$\lambda_0 = \left( \frac{9\pi}{8} \right)^{2/3} S^{2/3} - \frac{S}{\gamma - 1} . \quad (37)$$

Using this in (28), we find that

$$\mathcal{C}_1 \approx \frac{U_d - U_b}{s} - \left[ \frac{2}{S} - \left( \frac{9\pi}{8} \right)^{2/3} \frac{1}{S^{1/3}} + \frac{1}{\gamma - 1} \right] . \quad (38)$$

As our numerical analysis of (27) shows that the fraction  $x_0 \sim e^{-1/S}$  for small  $S$  (also, see Fig. 4), the variance (30) can be approximated by

$$\mathcal{C}_2 \approx \frac{U_d - U_b}{s} , \quad (39)$$

as also seen in (35) when  $s < U_d$ . The above equation also shows that the variance is larger than the mean as illustrated in Fig. 2. The above approximations are in good agreement with the exact results, see Fig. 2.

An analysis of the frequency distribution (27) for small  $S$  described in Appendix A7 suggests a Gaussian approximation for the frequency distribution,

$$x_k \approx \sqrt{\frac{1}{2\pi\mathcal{C}_2}} \exp \left[ -\frac{(k - \mathcal{C}_1)^2}{2\mathcal{C}_2} \right] , \quad (40)$$

where the mean and variance are given, respectively, by (38) and (39). Figure 4 compares the above approximation with the exact distribution (27) and we see a quite good agreement. However, it should be noted that unlike (40), the exact frequency distribution is not symmetric about the mean. The skewness defined as  $\mathcal{C}_3/\mathcal{C}_2^{3/2}$ , where  $\mathcal{C}_3$  is the third cumulant, is a measure of the asymmetry of the distribution. Due to (A16) in the stationary state, on neglecting the frequency  $x_0$ , we find a nonzero  $\mathcal{C}_3 = U/s$ . Figure 4 also shows that both mean and variance are considerably affected by beneficial mutations when  $s < U_b < U_d$ .

## 5. Dynamics of the population frequency

In the last section, we discussed the stationary state and now turn to the dynamics of the frequency distribution. We will focus on the time dependence of the average  $\mathcal{C}_1(t)$  starting from a monomorphic initial condition given by

$$x_k(0) = \delta_{k,k(0)} . \quad (41)$$

As the exact expression (26) for the time-dependent frequency involves a sum over a large number of eigenvalues, the dynamics are more efficiently studied by solving the differential equations (1a) and (1b) numerically. The results for the mean thus obtained are shown in Fig. 5 for two values of  $S$  and, we observe (i) a short time regime where the population is far from the stationary state, (ii) an intermediate time regime where the mean changes quickly and (iii) a long time relaxation regime where the population is close to the steady state.

The initially monomorphic population first spreads over the genotypic space due to mutations followed by an increase in the frequency of high fitness genotypes as a result of selection. At short enough times, one can understand the dynamics away from the stationary state by ignoring the boundary at  $k = 0$  (see, Sec. 5.1 below). Note that this holds even if the population is initially located in the zeroth fitness class since, as Fig. 5 shows, the mean initially increases because  $U_d > U_b$ . However once the population is close to the stationary state (i.e.,  $\mathcal{C}_1(t) - \mathcal{C}_1 \lesssim 1$  in Fig. 5), the boundary at the zeroth fitness class becomes important. As discussed in Sec. 5.2, at long enough times, it is sufficient to retain the minimum and second minimum eigenvalue in the sum over the eigenvalues in (26) to determine the time to relax to the stationary state.

### 5.1. Dynamics far from the stationary state

The dynamics of the  $n$ th cumulant  $\mathcal{C}_n(t)$  are described in Appendix A3. For the initial condition (41), at short enough times, we can set the frequency in the fittest class to be approximately zero in (A14) to obtain

$$\dot{\vec{\mathcal{C}}}(t) = -s\hat{D}\vec{\mathcal{C}}(t) + \vec{U} , \quad (42)$$

where  $\vec{\mathcal{C}}$  and  $\vec{U}$  are column vectors whose  $n$ th element is given by  $\mathcal{C}_n$  and  $U_d + (-1)^n U_b$  respectively and  $\hat{D}$  is an upper shift matrix with matrix element  $D_{mn} = \delta_{m+1,n}$ . The above equation can be straightforwardly solved for

arbitrary initial condition and for (41), we obtain

$$\mathcal{C}_n(t) = \frac{U_n}{s} \sinh(st) - \frac{U_{n+1}}{s} (\cosh(st) - 1) + k^{(0)} \delta_{n,1} . \quad (43)$$

Using (43) in (A13), the generating function of the population fraction,  $F(\xi, t) = \sum_{k=0}^{\infty} x_k(t) e^{-\xi k}$  can also be obtained and given by

$$\ln F(\xi, t) = -k^{(0)} \xi - \frac{U_d}{s} (1 - e^{-st})(1 - e^{-\xi}) - \frac{U_b}{s} (1 - e^{st})(e^{\xi} - 1) . \quad (44)$$

The above result generalises (53) of [14] who obtained it for special values of the parameters.

Due to (43), the approximate short time dynamics of mean  $\mathcal{C}_1(t)$  and variance  $\mathcal{C}_2(t)$  are given by

$$\mathcal{C}_1(t) = \frac{U_d}{s} (1 - e^{-st}) + \frac{U_b}{s} (1 - e^{st}) + k^{(0)} , \quad (45)$$

$$\mathcal{C}_2(t) = \frac{U_d}{s} (1 - e^{-st}) + \frac{U_b}{s} (e^{st} - 1) . \quad (46)$$

The above equations show that for  $t \ll 1/s$ , both mean and variance change linearly with time with a slope that depends only on the mutation rates but for longer times, the mean varies exponentially fast at a rate  $s^{-1}$ . The short time dynamics of the mean given by (45) are valid as long as  $|\mathcal{C}_1(t) - \mathcal{C}_1|$  is large and agree with the numerical results shown in Fig. 5.

### 5.2. Dynamics close to the stationary state

When beneficial mutations are absent, the frequency  $x_k(t)$  is Poisson-distributed with mean  $(U_d/s)(1 - e^{-st})$  [17, 7] and thus approaches the stationary state at rate  $s$ , independent of the deleterious mutation rate. When the beneficial mutation rate  $U_b$  is nonzero, the dynamical evolution of the frequency is given by (26). At large but finite times, it is a good approximation to retain only the terms containing the minimum and second minimum eigenvalues in the sum over the eigenvalues in (26). Thus the frequency  $x_k(t)$  relaxes to the steady state exponentially fast at rate

$$R = (\lambda_1 - \lambda_0) \sqrt{U_b U_d} , \quad (47)$$

where  $\lambda_1$  is the second minimum eigenvalue. Unlike  $\lambda_0$ , the second minimum eigenvalue  $\lambda_1$  is an increasing function of  $S$  as shown in Fig. 1.

The second minimum eigenvalue is calculated in Appendix A4 and A6 and given by

$$\lambda_1 = \begin{cases} \left(\frac{21\pi}{8}\right)^{2/3} S^{2/3} - \frac{S}{\gamma-1} & , S \ll 1 , \\ S + 2 & , S \gg 1 . \end{cases} \quad (48)$$

This yields the relaxation rate  $R = s + U_b$  for large  $S$  which shows that the population reaches the stationary distribution faster in the presence of beneficial mutations. For small  $S$ , we get  $R \sim s^{2/3}(U_b U_d)^{1/6}$  which approaches zero as  $s \rightarrow 0$  in accordance with the neutral case where the population never reaches a stationary state, see Appendix A1. The relaxation dynamics of the mean  $\mathcal{C}_1(t)$  are in agreement with the above results as shown in Fig. 5.

On comparing the results for short and long time dynamics, we find that while the former occurs over a time scale  $\sim s^{-1}$  independent of the mutation rates, the relaxation time is determined by both selection and mutation.

## 6. Conclusions

In this article, we presented the exact solution (26) for the frequency distribution at all times for the model defined by (1a) and (1b). Our results summarised in Table 1 generalise the earlier ones in [15, 10, 7] by including beneficial mutations and extend the treatment in [23, 14, 5, 21, 2, 19] to all times including the stationary state limit. We discussed the biologically realistic situation where the beneficial mutation rate is smaller than its deleterious counterpart [20] but, for completeness, we explore the parameter regime  $U_d \ll U_b$  in Appendix A8.

Here we considered a mutation scheme in which the mutation rate per sequence is same for all sequences, irrespective of their fitness. The evolution of an infinitely large population on additive fitness landscapes when the mutation rates depend linearly on the number of loci carrying deleterious allele has also been studied [26, 13] and the relationship of this mutation scheme with the one studied in this article is elucidated in Appendix A9. In the fitness-dependent mutation rate model [26, 13], when the number of loci carrying the deleterious allele is small, the beneficial mutation rate vanishes in the limit of infinitely long sequence. This has the immediate consequence that the stationary state properties are not affected by beneficial mutations in this mutation scheme [13].

In contrast, for the model studied here, the general effect of beneficial mutations is to decrease both mean and variance in the stationary state (see,

(28) and (30)) but the extent to which this happens depends on the strength of selection relative to mutations. We find that

(i) when  $U_b < U_d < s$ , beneficial mutations have a minor effect since the mean number of deleterious mutations in the absence of beneficial mutations is already close to zero,

(ii) when  $U_b < s < U_d$ , beneficial mutations decrease the mutational load significantly and the frequency is enhanced (diminished) in fitness classes below (above)  $U_d/s$  and

(iii) when  $s < U_b < U_d$ , both mean and variance decrease considerably and the frequency distribution is a nontrivial function.

Here we have focused on the deterministic evolution and ignored the effect of random genetic drift. However in a finite population, the bulk of the distribution is expected to be well described by deterministic distributions [5, 21, 2, 9]. A detailed study of the evolution of finite populations exploiting the results presented here will be taken up in future.

**Acknowledgements:** The authors thank Arul Lakshminarayan and three reviewers for several helpful comments.

## A1. Neutral dynamics using eigenfunction expansion method

The treatment below for the unnormalised frequency  $y_k(t)$  essentially follows Chap. 7, [24] and here we briefly describe our results. For  $S = 0$ , the solution of the eigenvalue equation (17b) is given by

$$\phi_k = C_+ a_+^k + C_- a_-^k, \quad k \geq 0, \quad (\text{A1})$$

where  $a_{\pm}$  are solutions of the quadratic equation  $a^2 + (\lambda - 2)a + 1 = 0$  and the coefficients  $C_+$  and  $C_-$  are related due to the boundary equation (17a). Since  $a_+ a_- = 1$ , it is convenient to write  $a_{\pm} = e^{\pm iq}$  where  $q$  is real. The latter condition is required to ensure that the eigenfunction  $\phi_k$  does not diverge at large  $k$  (see (3)). Using  $a_+ + a_- = 2 - \lambda$ , we find that the eigenvalues form a continuous spectrum and are given by

$$\lambda = 4 \sin^2 \left( \frac{q}{2} \right), \quad 0 \leq q \leq \pi. \quad (\text{A2})$$

Since the ratio  $C_+/C_-$  determined using (24) has unit modulus, we can write

$$\frac{C_+}{C_-} = -\frac{\gamma - 2 + e^{iq}}{\gamma - 2 + e^{-iq}} = e^{i2\eta(q)}, \quad (\text{A3})$$

and finally arrive at

$$\phi_k(q) = \sqrt{\frac{2}{\pi}} \cos(qk + \eta(q)), \quad (\text{A4})$$

where we have used the orthonormality condition (18) to determine the proportionality constant.

For the initial condition  $x_k(0) = \delta_{k,0}$ , on replacing the sums in (16) and (20) by integrals (as the eigenvalues are continuous), we obtain

$$\begin{aligned} y_k(t) &= \int_0^\pi \phi_k(q) \phi_0(q) e^{-2(1-\cos q)\sqrt{U_b U_d} t} dq, \\ &= \frac{2}{\pi} \int_0^\pi dq \sin q \frac{(\gamma - 2) \sin(qk) + \sin(qk + q)}{(\gamma - 2)^2 + 2(\gamma - 2) \cos q + 1} e^{-2(1-\cos q)\sqrt{U_b U_d} t} \end{aligned} \quad (\text{A5})$$

where the last expression follows on using (A3) in (A4). The above integral does not appear to be exactly solvable, but in the scaling limit  $q \rightarrow 0, k, t \rightarrow$



$\infty$  with  $q^2 t$  and  $qk$  finite, the above equation simplifies to give

$$y_k(t) \approx \frac{2}{\pi} \int_0^\infty dq \frac{(\gamma - 1)q \sin(qk) + q^2 \cos(qk)}{(\gamma - 1)^2} e^{-q^2 \sqrt{U_b U_d} t}, \quad (\text{A7})$$

$$= \frac{1}{2\sqrt{\pi}} \frac{k}{\gamma - 1} \frac{e^{-\frac{k^2}{4t\sqrt{U_b U_d}}}}{(t\sqrt{U_b U_d})^{3/2}}. \quad (\text{A8})$$

## A2. Some properties of the Bessel functions

If  $\mathcal{K}$  denotes  $J, Y$ , the Bessel function  $\mathcal{K}_\nu(z)$  is defined as the solution of the following differential equation (9.1.1,[1])

$$z^2 \frac{d^2 \mathcal{K}_\nu(z)}{dz^2} + z \frac{d\mathcal{K}_\nu(z)}{dz} + (z^2 - \nu^2) \mathcal{K}_\nu(z) = 0. \quad (\text{A9})$$

The Bessel function of the second kind  $Y_\nu(z)$  is related to the Bessel function of the first kind  $J_\nu(z)$  by (9.1.2,[1])

$$Y_\nu(z) = \frac{\cos(\nu\pi) J_\nu(z) - J_{-\nu}(z)}{\sin(\nu\pi)}. \quad (\text{A10})$$

For the series representation of  $J_\nu(z)$ , see (A17) below; the asymptotic expansions of  $J_\nu(z)$  for  $z > \nu$  and  $z < \nu$  are given in (A23) and (A29), respectively.

## A3. Cumulants of the number of deleterious mutations

Following [7], we first define the generating function of the frequency as

$$F(\xi, t) = \sum_{k=0}^{\infty} x_k(t) e^{-\xi k}. \quad (\text{A11})$$

Multiplying (1a) and (1b) by  $e^{-\xi k}$  and summing over  $k$ , we obtain

$$\frac{d \ln F(\xi, t)}{dt} = s \mathcal{C}_1(t) + s \frac{d \ln F(\xi, t)}{d\xi} - U_d(1 - e^{-\xi}) + U_b(e^\xi - 1) \left( 1 - \frac{x_0}{F(\xi, t)} \right). \quad (\text{A12})$$

The  $n$ th cumulant  $\mathcal{C}_n(t)$ ,  $n = 1, 2, \dots$  of the number of deleterious mutations is related to  $F(\xi, t)$  through

$$\ln F(\xi, t) = \sum_{n=1}^{\infty} \mathcal{C}_n(t) \frac{(-\xi)^n}{n!}. \quad (\text{A13})$$

Using the above equation in (A12), we get

$$\begin{aligned} \sum_{n=1}^{\infty} \dot{\mathcal{C}}_n(t) \frac{(-\xi)^n}{n!} &= -s \sum_{n=1}^{\infty} \mathcal{C}_{n+1}(t) \frac{(-\xi)^n}{n!} + U_d \sum_{n=1}^{\infty} \frac{(-\xi)^n}{n!} \\ &+ U_b \left( \sum_{n=1}^{\infty} \frac{\xi^n}{n!} \right) \left( 1 - \frac{x_0(t)}{F(\xi, t)} \right). \end{aligned} \quad (\text{A14})$$

Matching the coefficient of  $\xi^n$  for  $n = 1, 2$  on both sides, we find that

$$\dot{\mathcal{C}}_1(t) = -s\mathcal{C}_2(t) + U_d - U_b(1 - x_0(t)), \quad (\text{A15})$$

$$\dot{\mathcal{C}}_2(t) = -s\mathcal{C}_3(t) + U - U_b x_0(t)(1 + 2\mathcal{C}_1(t)). \quad (\text{A16})$$

#### A4. Approximate expressions for the eigenvalues when $S$ is large

The Bessel function of the first kind has the following series representation (9.1.10, [1]):

$$J_\nu(z) = \sum_{m=0}^{\infty} \frac{(-1)^m}{m!(\nu+m)!} \left( \frac{z}{2} \right)^{2m+\nu}. \quad (\text{A17})$$

For large  $S$ , keeping the first two terms in the above series and using it in the eigenvalue equation (24), we get

$$(\lambda - \gamma)(\lambda - S - 2) \approx 1. \quad (\text{A18})$$

Solving the above quadratic equation, we obtain the first two eigenvalues  $\lambda_0$  and  $\lambda_1$  given in (31) and (48) respectively. Figure 1 shows a comparison between the exact eigenvalues obtained numerically using (24) and the above approximations. Our numerical analysis of (24) also suggests that the  $\alpha$ th eigenvalue is given by

$$\lambda_\alpha \approx \alpha S + 2, \quad \alpha = 1, 2, \dots. \quad (\text{A19})$$

#### A5. Approximate expression for the stationary distribution when $S$ is large

Noting that the terms corresponding to  $m = 0, 1$  in the summand on the RHS of (33) contribute to the leading order term in  $U_b/s$ , we obtain

$$x_k \propto \frac{\mu^k}{k!} \left[ 1 + \frac{U_b}{s} \left( (\gamma_{EM} - H_k)(1 + \mu) - \frac{\mu}{k+1} \right) \right], \quad (\text{A20})$$

where  $\mu = U_d/s$ . In the above equation, we have used that  $k!/(k+\epsilon)! \approx 1 + \epsilon(\gamma_{EM} - H_k)$  for small  $\epsilon$  where  $\gamma_{EM} \approx 0.577\dots$  and  $H_k = \sum_{i=1}^k i^{-1}$  is the Harmonic number. On fixing the proportionality constant using (2), we recover (A7) of [12]:

$$x_k = \frac{e^{-\mu}\mu^k}{k!} \left[ 1 + \frac{U_b}{s} \left( \sum_{m=0}^{\infty} \frac{e^{-\mu}\mu^m}{m!} (H_m(1+\mu) + \frac{\mu}{m+1}) - H_k(1+\mu) - \frac{\mu}{k+1} \right) \right]. \quad (\text{A21})$$

For  $U_d < s$ , the above expression shows that the distribution is close to the Poisson distribution (5). However, for  $U_d > s$ , we obtain [12]

$$x_k \approx \frac{e^{-U_d/s}}{k!} \left( \frac{U_d}{s} \right)^k \left[ 1 + \frac{U_b U_d}{s^2} \ln \left( \frac{U_d}{sk} \right) \right], \quad (\text{A22})$$

on using  $\sum_{m=0}^{\infty} z^m H_m/m! \approx e^z \ln z$  for large  $z$  and  $H_k \approx \ln k + \gamma_{EM}$  for large  $k$  in (A21). This result shows that the beneficial mutations enhance the frequency in fitness classes  $k < U_d/s$  and diminish it in higher ones (also, see Fig. 3).

## A6. Approximate expressions for the eigenvalues when $S$ is small

For the Bessel function  $J_\nu(z)$ ,  $z > \nu$ , the asymptotic expansion for large orders is given by (9.3.3, [1])

$$J_\nu(\nu \sec \beta) \sim \frac{\cos(\nu(\tan \beta - \beta) - (\pi/4))}{\sqrt{(\nu/2)\pi \tan \beta}}, \quad 0 < \beta < \pi/2. \quad (\text{A23})$$

As shown in Fig. 1, the eigenvalues  $\lambda_0, \lambda_1$  are an increasing function of  $S$  and approach zero as  $S \rightarrow 0$  (also, see (A2) for the neutral case). Then using (A23) in (24) and carrying out a small  $\lambda$  expansion, we obtain

$$\frac{\cos(\frac{2\lambda^{3/2}}{3S} - \frac{\pi}{4} - \sqrt{\lambda})}{\cos(\frac{2\lambda^{3/2}}{3S} - \frac{\pi}{4})} = \gamma - \lambda. \quad (\text{A24})$$

After some algebra, the above simplifies to

$$\tan \left( \frac{2\lambda^{3/2}}{3S} \right) = -\frac{\gamma - 1 + \sqrt{\lambda} - \lambda}{\gamma - 1 - \sqrt{\lambda} - \lambda}. \quad (\text{A25})$$

The above equation immediately suggests that the eigenvalue  $\lambda \sim S^{2/3}$  so that the RHS can be nonzero and finite. Guided by this observation and a numerical analysis of (24), we expect that the  $\alpha$ th eigenvalue is of the following form:

$$\lambda_\alpha = \lambda_\alpha^{(0)} S^{2/3} + \lambda_\alpha^{(1)} S. \quad (\text{A26})$$

Substituting this in (A25) and expanding both sides of the equation for small  $S$ , we find that

$$\lambda_\alpha^{(0)} = \left( \frac{3\pi(4\alpha + 3)}{8} \right)^{2/3}, \quad (\text{A27})$$

$$\lambda_\alpha^{(1)} = -\frac{1}{\gamma - 1}. \quad (\text{A28})$$

As we have assumed  $\lambda$  to be small to arrive at (A24), the above results for the eigenvalues are valid for small  $\alpha$ . For larger  $\alpha$ , our numerical analysis of (24) suggests that the eigenvalues increase linearly with  $S$ .

#### A7. Approximate expression for the stationary distribution when $S$ is small

The asymptotic expansion of the Bessel function  $J_\nu(z)$ ,  $z < \nu$  for large orders is given by (9.3.1, [1])

$$J_\nu(z) \sim \frac{1}{\sqrt{2\pi\nu}} \left( \frac{ze}{2\nu} \right)^\nu. \quad (\text{A29})$$

Using this in (27), we obtain the steady state frequency to be

$$x_k \propto \left( \sqrt{\frac{U_d}{U_b}} \right)^{k+\delta} \frac{1}{\sqrt{k+\delta}} \left( \frac{e/S}{k+\delta} \right)^{k+\delta}, \quad (\text{A30})$$

where  $\delta = (2 - \lambda_0)/S$  and  $\lambda_0$  is given by (37). A Gaussian approximation for the above expression can be obtained by writing  $x_k \propto e^{I(k)}$  and expanding  $I(k)$  about its turning point  $\tilde{k} = (U_d/s) - \delta$  up to quadratic orders in the deviation  $k - \tilde{k}$ . On fixing the normalisation, we obtain

$$x_k \approx \sqrt{\frac{s}{2\pi U_d}} \exp \left[ -\frac{s}{2U_d} \left( k - \frac{U_d}{s} + \frac{2 - \lambda_0}{S} \right)^2 \right]. \quad (\text{A31})$$

The mean and variance of the above distribution differs from (38) and (39) by a factor  $U_b/s$  and is therefore a good approximation when  $U_b \ll U_d$ .

### A8. Stationary state distribution when the deleterious mutation rate is smaller than the beneficial one

For completeness, here we consider the parameter regime in which  $U_d < U_b$  within a perturbation theory in  $U_d$ . We begin by expanding the steady state fraction in a power series in  $U_d$  as

$$x_k = \sum_{n=0}^{\infty} U_d^n \frac{x_k^{(n)}}{n!} , \quad (\text{A32})$$

where  $x_k^{(n)}$  is the  $n$ th derivative of  $x_k$  with respect to  $U_d$  evaluated at  $U_d = 0$ . When the deleterious mutation rate is zero, as the entire population is in the zeroth fitness class, we have  $x_k^{(0)} = \delta_{k,0}$ . As a result, the mean  $\mathcal{C}_1^{(0)} = \sum_{k=0}^{\infty} k x_k^{(0)} = 0$ . Using (A32) in (1a) and (1b) in the steady state and retaining terms to leading order in  $U_d$ , we obtain

$$s\mathcal{C}_1^{(1)} = 1 - U_b x_1^{(1)} \quad (\text{A33})$$

$$U_b x_2^{(1)} = (U_b + s)x_1^{(1)} - 1 \quad (\text{A34})$$

$$U_b x_k^{(1)} = (U_b + s(k-1))x_{k-1}^{(1)} , \quad k \geq 3 . \quad (\text{A35})$$

The last equation implies that the steady state fraction is a monotonically increasing function of  $k$ ; however, since each frequency is bounded above by unity and the total fraction must also add up to one, to obtain a sensible result to linear order in  $U_d$ , the fraction  $x_k^{(1)}$  must be zero for all  $k \geq 2$ . This immediately yields

$$x_0 = 1 - \frac{U_d}{U_b + s} + \mathcal{O}(U_d^2) \quad (\text{A36})$$

$$x_1 = \frac{U_d}{U_b + s} + \mathcal{O}(U_d^2) \quad (\text{A37})$$

$$x_k = \mathcal{O}(U_d^2) , \quad k \geq 2 . \quad (\text{A38})$$

Thus the fraction in the zeroth fitness class decreases linearly with  $U_d$  when the deleterious mutation rate is smaller than the beneficial one; in contrast, the fraction  $x_0$  decays exponentially or faster with  $U_d$  when  $U_d > U_b$  (see Sec. 4.1 and 4.2).

## A9. Comparison of mutation schemes

In an infinitely large population of finite diallelic sequences of length  $L$  in which the wild type allele mutates with rate  $\mu$  and the back mutation occurs with rate  $\nu$ , the frequency  $x_k(t)$  of a sequence with  $k$  deleterious mutations and fitness  $w_k = -sk$  evolves in continuous time as [26, 13]

$$\dot{x}_k = (k+1)\nu x_{k+1} + (L-k+1)\mu x_{k-1} - [(L-k)\mu + k\nu]x_k - s(k-\bar{k})x_k, \quad (\text{A39})$$

where  $x_{-1} = x_{L+1} = 0$ . In the limit  $\mu, \nu \rightarrow 0$  and  $L \rightarrow \infty$ , one can define the deleterious and beneficial mutation rate per sequence as  $U_d = L\mu$  and  $U_b = L\nu$ , and rewrite the above equation as

$$\dot{x}_k = \epsilon_{k+1}U_b x_{k+1} + (1-\epsilon_{k-1})U_d x_{k-1} - [(1-\epsilon_k)U_d + \epsilon_k U_b]x_k - s(k-\bar{k})x_k, \quad (\text{A40})$$

where  $\epsilon_k = k/L$ . In a well adapted population in which the number of loci carrying the deleterious allele is small, the back mutations to the wild type allele can be ignored. More precisely, when the number of deleterious mutations scales sublinearly with  $L$ , the fraction  $\epsilon_k \rightarrow 0$  for an infinitely long sequence and we obtain the model defined by (1a) and (1b) with  $U_b = 0$ . Similarly, in a maladapted population in which  $\epsilon_k \rightarrow 1$ , the model (A40) reduces to the one studied in this article with  $U_d = 0$ . The model defined by (1a) and (1b) thus interpolates between the two limits of the model (A40) described above.

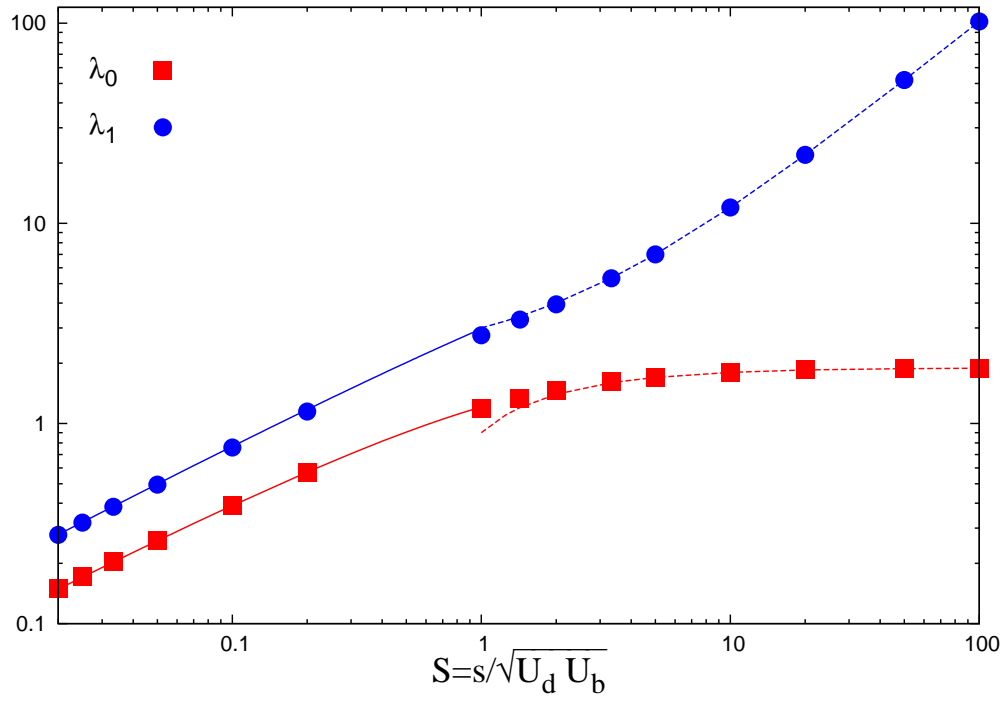


Figure 1: Variation of the minimum and second minimum eigenvalue  $\lambda_0$  and  $\lambda_1$  with  $S$ . The points are obtained by solving (24) numerically and the lines show the approximate expressions (37) and (31) for  $\lambda_0$  and (48) for  $\lambda_1$  for  $U_b = 0.01 U_d, s = 0.001$ .

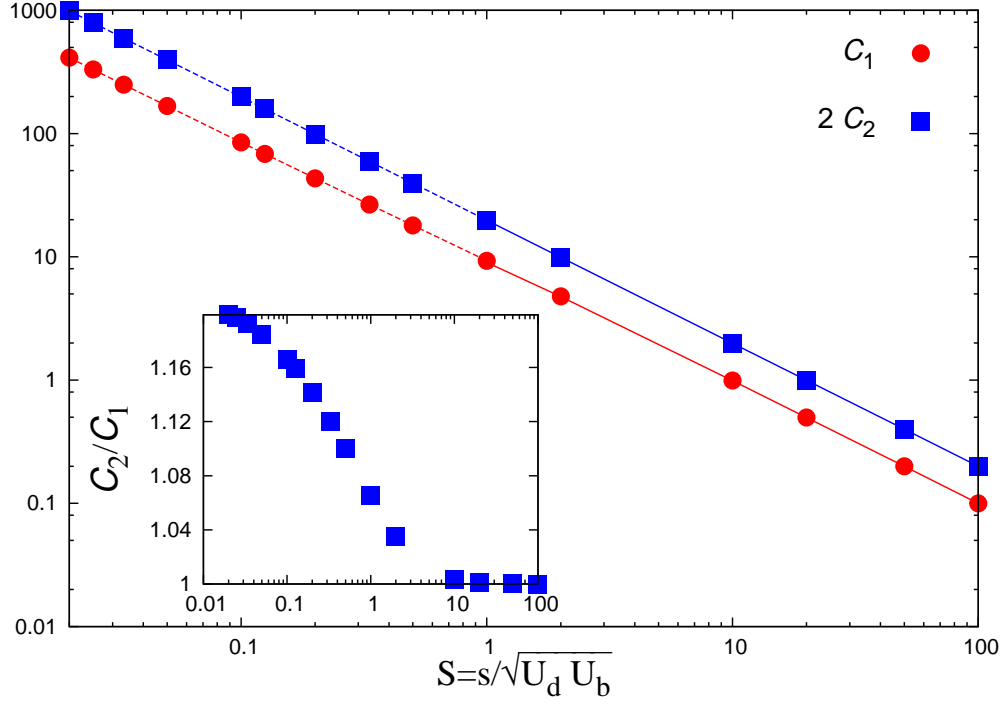


Figure 2: Main figure shows the exact mean  $\mathcal{C}_1$  and variance  $\mathcal{C}_2$  in the stationary state given, respectively, by (28) and (30) for  $U_b = 0.01 U_d, s = 0.001$ . The lines show the approximate expressions (34), (38) for mean and (35), (39) for variance. The variance is scaled by a factor 2 for clarity. The exact ratio  $\mathcal{C}_2/\mathcal{C}_1$  shown in the inset supports the non-Poissonian nature of the frequency distribution in the steady state.



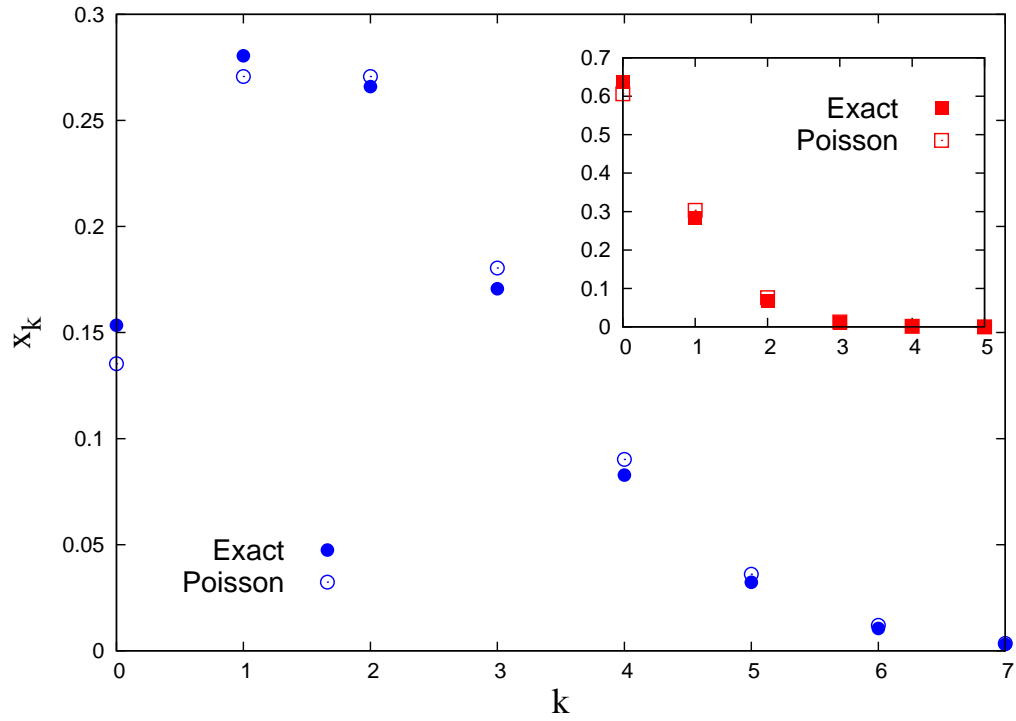


Figure 3: Steady state distribution when  $s \gg \sqrt{U_b U_d}$ : Comparison of the exact distribution (27) and Poisson distribution (5) for  $U_b = 0.005, s = 0.1, U_d = 0.2$  (main) and  $U_b = 0.01, U_d = 0.05, s = 0.1$  (inset).

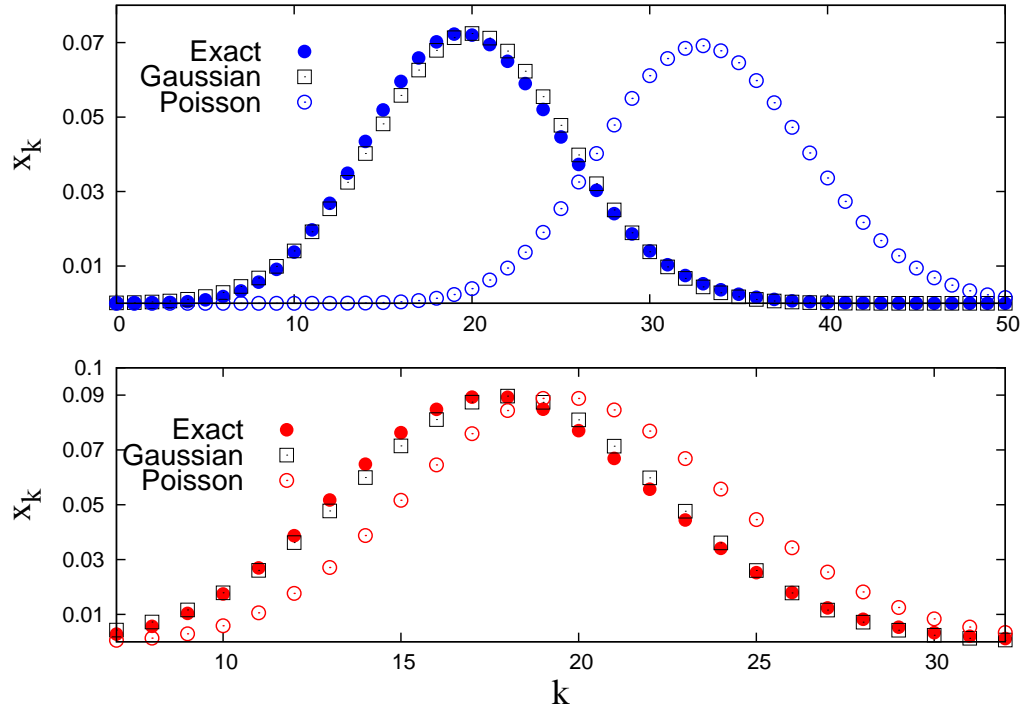


Figure 4: Steady state distribution when  $s \ll \sqrt{U_b U_d}$ : Comparison of the exact distribution (27), Gaussian approximation (40) and Poisson distribution (5). The parameters in the top and bottom panel are  $s = 0.003, U_b = 0.009, U_d = 0.1$  and  $U_b = 0.001, s = 0.005, U_d = 0.1$ , respectively.

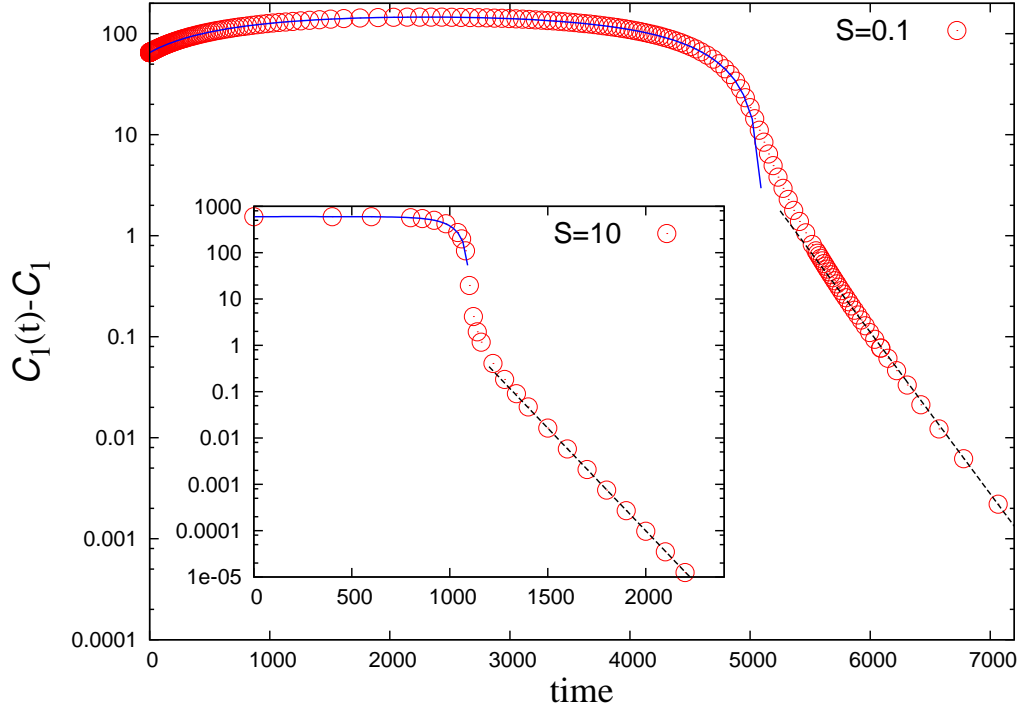


Figure 5: Dynamics of the mean deviation from the stationary state,  $\mathcal{C}_1(t) - \mathcal{C}_1$  for two values of  $S = s/\sqrt{U_b U_d}$  with initial population located in the fitness class  $k^{(0)}$ . The exact dynamics obtained by numerically integrating (1a) and (1b) are shown by points while the solid (blue) lines show the short time dynamics (45) and the broken (black) line shows the relaxation dynamics  $b e^{-Rt}$ , where the relaxation rate  $R$  is given by (47). The parameters in the main and inset are  $U_b = s = 0.001, U_d = 0.1, k^{(0)} = 150, b = 4.515 \times 10^8, \mathcal{C}_1 = 84.8958$  and  $U_b = 0.0001, s = U_d = 0.01, k^{(0)} = 600, b = 71561, \mathcal{C}_1 = 0.990147$ , respectively.

## References

- [1] M. Abramowitz and I.A. Stegun. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. Dover, 1964.
- [2] E. Brunet, I. M. Rouzine and C. O. Wilke. The Stochastic Edge in Adaptive Evolution. *Genetics*, 179: 603–620, 2008.
- [3] M. Bulmer. The selection-mutation-drift theory of synonymous codon usage. *Genetics*, 149:897–907, 1991.
- [4] R. Courant and D. Hilbert. *Methods of Mathematical Physics, Volume 1*. Dover, 1953.
- [5] M.M. Desai and D.S. Fisher. Beneficial mutation-selection balance and the effect of linkage on positive selection. *Genetics*, 176:1759–1798, 2007.
- [6] M. Ehrhardt and R. E. Mickens. Solutions to the discrete Airy equation: application to parabolic equation calculations. *J. Comp. Appl. Math.*, 172:183–206, 2004.
- [7] A. Etheridge, P. Pfaffelhuber, and A. Wakolbinger. How often does the ratchet click? facts, heuristics, asymptotics. In J. Blath, P. Mörters, and M. Scheutzow, editors, *Trends in Stochastic Analysis, London Mathematical Society Lecture Note Series 353*, pages 365–390. Cambridge University Press, 2009.
- [8] S. Flügge. *Practical quantum mechanics*. Springer-Verlag, 1974.
- [9] S. Goyal, D. J. Balick, E. R. Jerison, R. A. Neher, B. I. Shraiman, and M. M. Desai. Dynamic mutation selection balance as an evolutionary attractor. *Genetics*, 191:1309–1319, 2012.
- [10] J. Haigh. The accumulation of deleterious genes in a population - Muller’s ratchet. *Theoret. Population Biol.*, 14:251–267, 1978.
- [11] K. Jain and S. Seetharaman. Nonlinear deterministic equations in biological evolution. *J. Nonlin. Math. Phys.*, 18:321–338, 2011.
- [12] A. James and K. Jain. Fixation probability of rare nonmutator and evolution of mutation rates. *Ecology and Evolution*, 6:755–764, 2016.

- [13] S. John and K. Jain. Effect of drift, selection and recombination on the equilibrium frequency of deleterious mutations. *J. theo. Biol.*, 365:238–246, 2015.
- [14] D. A. Kessler, H. Levine, D. Ridgway, and L. Tsimring. Evolution on a smooth landscape. *J. Stat. Phys.*, 87:519 – 544, 1997.
- [15] M. Kimura and T. Maruyama. The mutational load with epistatic gene interactions in fitness. *Genetics*, 54:1337–1351, 1966.
- [16] W.-H. Li. Models of nearly neutral mutations with particular implications for nonrandom usage of synonymous codons. *J Mol Evol*, 24:337–345, 1987.
- [17] L. P. Maia, D. F. Botelho, and J. F. Fontanari. Analytical solution of the evolution dynamics on a multiplicative-fitness landscape. *J. Math. Biol.*, 47:453–456, 2003.
- [18] G. A. T. McVean and B. Charlesworth. A population genetic model for the evolution of synonymous codon usage: patterns and predictions. *Genet. Res.*, 74:145–158, 1999.
- [19] S.-C. Park, D. Simon, and J. Krug. The speed of evolution in large asexual populations. *J. Stat. Phys.*, 138:381–410, 2010.
- [20] L. Perfeito, L. Fernandes, C. Mota, and I. Gordo. Adaptive mutations in bacteria: high rate and small effects. *Science*, 317:813–815, 2007.
- [21] I.M. Rouzine, E. Brunet, and C.O. Wilke. The traveling-wave approach to asexual evolution: Muller’s ratchet and speed of adaptation. *Theo. Pop. Biol.*, 73:24–46, 2008.
- [22] C.J. Thompson and J.L. McBride. On Eigen’s theory of the self-organization of matter and the evolution of biological macromolecules. *Math. Biosci.*, 21:127, 1974.
- [23] L. S. Tsimring, H. Levine, and D. A. Kessler. RNA virus evolution via a fitness-space model. *Phys. Rev. Lett.*, 76:4440–4443, 1996.
- [24] N. G. van Kampen. *Stochastic processes in physics and chemistry*. North Holland Personal Library, 1997.

- [25] S. Wielgoss, J.E. Barrick, O. Tenaillon, M.J. Wiser, W.J. Dittmar, S. Cruveiller, B. Chane-Woon-Ming, C. Médigue, R. E. Lenski, and D. Schneider. Mutation rate dynamics in a bacterial population reflect tension between adaptation and genetic load. *Proc. Natl. Acad. Sci USA*, 110:222–227, 2013.
- [26] G. Woodcock and P. G. Higgs. Population evolution on a multiplicative single-peak fitness landscape. *J. theor. Biol.*, 179:61–73, 1996.